SEQUENCE LISTING

<110> DELBAC, FREDERIC
DANCHIN, ANTOINE
VIVARES, CHRISTIAN

<120> MICROSPORIDIAN POLAR TUBE PROTEINS, NUCLEIC ACIDS CODING FOR THESE PROTEINS AND THEIR APPLICATIONS

<130> 1566-00

<140> 09/755,456
<141> 2001-01-05

<150> PCT/FR99/01630
<151> 1999-07-06

<150> FR 98/08692
<151> 1998-07-07

<160> 10

<170> PatentIn Ver. 2.1

<210> 1
<211> 1830
<212> DNA
<213> Encephalitozoon cuniculi

<220>
<221> CDS
<222> (345)..(1529)

<400> 1
gaattcagat gcctcatacc ttgggattaa aaaattgatg ttcatttgtt atatatcctg 60

ggcggacagg ccggctcgta ttcttcaggg gtgtcgccta cccagtgcac aggaggttcc 120

ggaggtgtct tggatggaaa gtaaggccat ttgtgggttc tcatccatgt catcgtccct 180

ttcggctgtt tcaccaagat ccaattattc ctccaggact ttcaaccctc agaatggaaa 240

cagagatgaa actctctgtg caaatcgtag atatcgattg gagacattga aaccacggag 300

tttgaaataa aagtataaat acctccgaaa acgcagagtt taag atg aaa ggt att 356
Met Lys Gly Ile
1

tct aag atc ctc tct gcc tct att gcc ctg atg aag ttg gag aat gtc 404
Ser Lys Ile Leu Ser Ala Ser Ile Ala Leu Met Lys Leu Glu Asn Val
5 10 15 20

tat tca gca acc gca ctg tgc agc aat gca tat ggc cta act ccg gga 452
Tyr Ser Ala Thr Ala Leu Cys Ser Asn Ala Tyr Gly Leu Thr Pro Gly
25 30 35

```
caa cag ggt atg gct cag cag ccg tcg tat gtg ctg atc ccc agc acc   500
Gln Gln Gly Met Ala Gln Gln Pro Ser Tyr Val Leu Ile Pro Ser Thr
            40                  45                  50

ccg gga acc ata gca aac tgt gca agc ggt tca cag gac aca tat tct   548
Pro Gly Thr Ile Ala Asn Cys Ala Ser Gly Ser Gln Asp Thr Tyr Ser
        55                  60                  65

cct tct ccc gct gca ccc aca tct cca gtg act ccg ggg aaa act agc   596
Pro Ser Pro Ala Ala Pro Thr Ser Pro Val Thr Pro Gly Lys Thr Ser
    70                  75                  80

gag aat gag aca tct cca tcg gct cct gca gaa gat gta gga aca tgc   644
Glu Asn Glu Thr Ser Pro Ser Ala Pro Ala Glu Asp Val Gly Thr Cys
85                  90                  95                  100

aag att gcc gta ttg aag cac tgc gac gca cca gga aca aca tca ggg   692
Lys Ile Ala Val Leu Lys His Cys Asp Ala Pro Gly Thr Thr Ser Gly
                105                 110                 115

acg aca cca ggg tca ggg cct tgt gaa acc cca gag cag caa cag cct   740
Thr Thr Pro Gly Ser Gly Pro Cys Glu Thr Pro Glu Gln Gln Gln Pro
            120                 125                 130

ttg tca gtg atc tcc acc act cct gcc gta ccg gtg act gtg gag tct   788
Leu Ser Val Ile Ser Thr Thr Pro Ala Val Pro Val Thr Val Glu Ser
        135                 140                 145

gca cag tct cca tct gtt gtg cca gtt gtt cct gtc gtt gct cac cac   836
Ala Gln Ser Pro Ser Val Val Pro Val Val Pro Val Val Ala His His
    150                 155                 160

cag gca gtt cca ggc tac tac aac aat gga aca tcc ggt att cct gga   884
Gln Ala Val Pro Gly Tyr Tyr Asn Asn Gly Thr Ser Gly Ile Pro Gly
165                 170                 175                 180

cag caa cag atc ctt tct ggc act ctt ccc cca gga gcc act ttg tgt   932
Gln Gln Gln Ile Leu Ser Gly Thr Leu Pro Pro Gly Ala Thr Leu Cys
                185                 190                 195

cag gga cag gcc atg cct agc act cct gga cag caa cag atc ctt tct   980
Gln Gly Gln Ala Met Pro Ser Thr Pro Gly Gln Gln Gln Ile Leu Ser
            200                 205                 210

ggc act ctt ccc cca ggg gtc act ttg tgt cag gga cag gcc acg cct  1028
Gly Thr Leu Pro Pro Gly Val Thr Leu Cys Gln Gly Gln Ala Thr Pro
        215                 220                 225

agc act cct ggg cag caa cag gtc ctt tct ggc act ctt ccc cca gga  1076
Ser Thr Pro Gly Gln Gln Gln Val Leu Ser Gly Thr Leu Pro Pro Gly
    230                 235                 240

gtc act ttg tgt cag gga cag gcc acg cct agc act cct ggg cag caa  1124
Val Thr Leu Cys Gln Gly Gln Ala Thr Pro Ser Thr Pro Gly Gln Gln
245                 250                 255                 260
```

```
cag gtc ctt tct ggc acc ctt ctc cca gga gcc act ttg tgt cag gat   1172
Gln Val Leu Ser Gly Thr Leu Leu Pro Gly Ala Thr Leu Cys Gln Asp
                265                 270                 275

caa ggt atg cct gga aca tcc gga gtt cct gga cag cag gga cag tct   1220
Gln Gly Met Pro Gly Thr Ser Gly Val Pro Gly Gln Gln Gly Gln Ser
            280                 285                 290

agt gga cag tgt tgt gcc cct cag att cca aac cct gtc atg ccg cca   1268
Ser Gly Gln Cys Cys Ala Pro Gln Ile Pro Asn Pro Val Met Pro Pro
        295                 300                 305

tcc atg aac att agt gga aat ggg tat cct tct tct acc gca tac agc   1316
Ser Met Asn Ile Ser Gly Asn Gly Tyr Pro Ser Ser Thr Ala Tyr Ser
    310                 315                 320

cca aac ctc gga tca ctg gga tcc tgt gtt gac ata cag aag acg ggg   1364
Pro Asn Leu Gly Ser Leu Gly Ser Cys Val Asp Ile Gln Lys Thr Gly
325                 330                 335                 340

ggg aca tcc tgc gag caa aaa ccc gag aag tcc gcc acg cag tat gcc   1412
Gly Thr Ser Cys Glu Gln Lys Pro Glu Lys Ser Ala Thr Gln Tyr Ala
                345                 350                 355

atg gag gcc tgt gca aca cca aca cca acg gtt att ata ggc aac agc   1460
Met Glu Ala Cys Ala Thr Pro Thr Pro Thr Val Ile Ile Gly Asn Ser
            360                 365                 370

gag tat ctt gtt gga cca gga atg tac aat gca att aac tct cca tgc   1508
Glu Tyr Leu Val Gly Pro Gly Met Tyr Asn Ala Ile Asn Ser Pro Cys
        375                 380                 385

aac act gct gtc caa tgc tgc taggctaaaa taaaacgagt ttaatcttct      1559
Asn Thr Ala Val Gln Cys Cys
    390                 395

ttttcttcgg tcttttggaa cgttggatgg ggatggagga gtctatgggc tgaagtgaaa 1619

tgccaacact tcttctgccc aagaacacat tcggatgttc ttcctgtggc caggagtttg 1679

gtaacaggat tccccgagga tttagcagcc ttggagtacc atgattgaat cagtattaaa 1739

cttctcaaat tattttattc tttctgtttt atatcccgag ccaatctgag aagaatgcct 1799

cgaattcaag ctcccttaga agtgtgggat c                                 1830
```

<210> 2
<211> 1740
<212> DNA
<213> Encephalitozoon cuniculi

<220>
<221> CDS
<222> (458)..(1288)

<400> 2
aagcttctga acaagcgcta accctctttc agaatatata aagcaatcca tacaacttct 60

```
ccatccatcc cggtgctgtt tctttggagg caaaacagag gaggtggcga tatcgatggt 120

gcatccataa tatatacaag acactccagg ctgcaactga atcaacacac tccatcccct 180

caggaagtcg gtaaacttgc cttgaaaata gccaatggat gtctccaggc tttataccat 240

gcacagctat atcttggcct gaagtgcact ttcaggtggg gctttgttac attgcggtgt 300

tttggattac ctgatataat ttgttaccca ctgagtcaag tcgaaaccag tagtccgcag 360

atttctaaca gagaggaaag actggaggta atttgtggct tttgaaacat gcacagcaaa 420

ataaaatata aaagaagcct tttgcacact accaaag atg ttg tta ctt ctc gcc  475
                                          Met Leu Leu Leu Leu Ala
                                            1               5

ata act gct gtt gtt agc gcc acg atg gtc cat cct tca gct gtt gtt   523
Ile Thr Ala Val Val Ser Ala Thr Met Val His Pro Ser Ala Val Val
            10                  15                  20

cca cag ccc gca gca cct ctc cat gtc gtt ccc cca cag cag caa atg   571
Pro Gln Pro Ala Ala Pro Leu His Val Val Pro Pro Gln Gln Gln Met
        25                  30                  35

ggc atg gtt aac gga tgc acc agc aag aaa cta gag ggt gca gaa ata   619
Gly Met Val Asn Gly Cys Thr Ser Lys Lys Leu Glu Gly Ala Glu Ile
    40                  45                  50

atg aga agg aac atg att gag tgc cag aaa aga agc tcg gag gca aca   667
Met Arg Arg Asn Met Ile Glu Cys Gln Lys Arg Ser Ser Glu Ala Thr
55                  60                  65                  70

aag gcg atg att gaa agg gca aat gaa aag gct gta gaa tca ttc aac   715
Lys Ala Met Ile Glu Arg Ala Asn Glu Lys Ala Val Glu Ser Phe Asn
                75                  80                  85

aag gaa gtt agc aaa gga cct agc caa aag gat gga ggc cag tgc ata   763
Lys Glu Val Ser Lys Gly Pro Ser Gln Lys Asp Gly Gly Gln Cys Ile
            90                  95                  100

gaa aaa gct gta caa ggt acc gat agg tgt att ctc gct gga ata atc   811
Glu Lys Ala Val Gln Gly Thr Asp Arg Cys Ile Leu Ala Gly Ile Ile
        105                 110                 115

gat aag gcg gtg aac aag cgc aag tac aga atc tca gat gtg gag aac   859
Asp Lys Ala Val Asn Lys Arg Lys Tyr Arg Ile Ser Asp Val Glu Asn
    120                 125                 130

agc acc tcg ctc tac aga gga gac aag cta att gcc cta att gtc aat   907
Ser Thr Ser Leu Tyr Arg Gly Asp Lys Leu Ile Ala Leu Ile Val Asn
135                 140                 145                 150

gtc gac tat ggg ctg cag ccg atc act aag cca aag aag aag aag tcc   955
Val Asp Tyr Gly Leu Gln Pro Ile Thr Lys Pro Lys Lys Lys Lys Ser
                155                 160                 165
```

```
aag ata atg gcg aat ctc cct cag ccg aag aga gag atg tat ttc aac   1003
Lys Ile Met Ala Asn Leu Pro Gln Pro Lys Arg Glu Met Tyr Phe Asn
            170                 175                 180

caa atc ggt cag ctt gtt gga gca aga gga acg ttc ccc cag gaa aac   1051
Gln Ile Gly Gln Leu Val Gly Ala Arg Gly Thr Phe Pro Gln Glu Asn
        185                 190                 195

aag gag gac tgc aag cct tgt gag ggt ccc aag aag act gtt gaa act   1099
Lys Glu Asp Cys Lys Pro Cys Glu Gly Pro Lys Lys Thr Val Glu Thr
    200                 205                 210

act tct gag aaa tgt aat ctt ggg tgc gag ctt aaa gga aca tct gct   1147
Thr Ser Glu Lys Cys Asn Leu Gly Cys Glu Leu Lys Gly Thr Ser Ala
215                 220                 225                 230

ctg ata agc aag gcc ata cag aag aag gaa gtc aag gac acg aag gaa   1195
Leu Ile Ser Lys Ala Ile Gln Lys Lys Glu Val Lys Asp Thr Lys Glu
                235                 240                 245

ggg gag aaa agt gca agc cag gac tct gat ggc gag ggc act gct gag   1243
Gly Glu Lys Ser Ala Ser Gln Asp Ser Asp Gly Glu Gly Thr Ala Glu
            250                 255                 260

gat gcg gaa gta cag caa cct tct gcg gac ggc gag ggt cta gag       1288
Asp Ala Glu Val Gln Gln Pro Ser Ala Asp Gly Glu Gly Leu Glu
        265                 270                 275

taatttttaa attaaaatct ccctggattg aatcttcaag tgcttttgtg aaagactttg 1348

ggaacatttc gtgaaggcta acataaattg ttaatctcag gtcactcgat ggaatagtca 1408

attcgtattt cctttccttg gatggtctgc cccaccagcc tgttcctggc agttatcgca 1468

tcgtcgacag agtcaaactg aacgaatcca tatcctttgg acatcttctt gtattggtcg 1528

tagactatta ctacccgata gttcagtatc tcactgatcc tctccttgag aaggtctcta 1588

acgtcgtctt cggttatgtg tgctcccagc ccaaatatcc ctatcgccct ggagggagac 1648

ccgtttctct ttgctttaag tgcatatctt tcgttttat aggagcttgg atctgttcct 1708

tcgtatcccc ttgtcgggcg ctccacctcg ag                               1740
```

<210> 3
<211> 1116
<212> DNA
<213> Encephalitozoon intestinalis

<220>
<221> CDS
<222> (1)..(1113)

<400> 3

```
atg aaa ggt att tct aag gtt ctc tca gcc tct att gtc cta atg aag   48
Met Lys Gly Ile Ser Lys Val Leu Ser Ala Ser Ile Val Leu Met Lys
  1               5                  10                  15
```

```
ttg aag ggt gtc tat tct aca act gtg ctg tgt gga gat tca aca caa   96
Leu Lys Gly Val Tyr Ser Thr Thr Val Leu Cys Gly Asp Ser Thr Gln
            20                  25                  30

gga ctg cag ggc aca acc caa ccg tca tat gtg ctg gtt cct agt gca  144
Gly Leu Gln Gly Thr Thr Gln Pro Ser Tyr Val Leu Val Pro Ser Ala
        35                  40                  45

cca gag aca ata gcc aac tgt gga tac agt cca cag aac atg tat gtc  192
Pro Glu Thr Ile Ala Asn Cys Gly Tyr Ser Pro Gln Asn Met Tyr Val
    50                  55                  60

cct tct act cct act acc atg cct tcc aca gtg cca ggc aca act ggt  240
Pro Ser Thr Pro Thr Thr Met Pro Ser Thr Val Pro Gly Thr Thr Gly
65                  70                  75                  80

gag agc gag aca cct act tct cca aca tca tct cct aca gag gat gtg  288
Glu Ser Glu Thr Pro Thr Ser Pro Thr Ser Ser Pro Thr Glu Asp Val
                85                  90                  95

gga aca tgc aag att gct gtt gta aag cat tgt gat gca cca gga aca  336
Gly Thr Cys Lys Ile Ala Val Val Lys His Cys Asp Ala Pro Gly Thr
            100                 105                 110

tca tca aca cct tgc gaa ccg gaa cag act ttg gcc ccc tct cag cca  384
Ser Ser Thr Pro Cys Glu Pro Glu Gln Thr Leu Ala Pro Ser Gln Pro
        115                 120                 125

gta gca gct aca att gcc aca cca ctg gtt gtt gct tct gtg cag acg  432
Val Ala Ala Thr Ile Ala Thr Pro Leu Val Val Ala Ser Val Gln Thr
    130                 135                 140

ccg caa gca gct gtt acc atc ctt act cca aag gcc gtc tct gcc cag  480
Pro Gln Ala Ala Val Thr Ile Leu Thr Pro Lys Ala Val Ser Ala Gln
145                 150                 155                 160

ccg gca acc atc att tct cca ttc aac cag gca cca ggc tac tac aat  528
Pro Ala Thr Ile Ile Ser Pro Phe Asn Gln Ala Pro Gly Tyr Tyr Asn
                165                 170                 175

agt gca att ccc ggg caa ata ctt aca ggt aat gtt ctc tct cca agt  576
Ser Ala Ile Pro Gly Gln Ile Leu Thr Gly Asn Val Leu Ser Pro Ser
            180                 185                 190

gcc tct tct tgc caa gtg gtg ccc gga aca aca gga agc tcc acc ccc  624
Ala Ser Ser Cys Gln Val Val Pro Gly Thr Thr Gly Ser Ser Thr Pro
        195                 200                 205

cag cag cta cca ggc gct gtt tca tct gga acc att cct tgc caa ata  672
Gln Gln Leu Pro Gly Ala Val Ser Ser Gly Thr Ile Pro Cys Gln Ile
    210                 215                 220

gta cag gga act caa agt agc gga aac acc cct gga cag caa ttc ttg  720
Val Gln Gly Thr Gln Ser Ser Gly Asn Thr Pro Gly Gln Gln Phe Leu
225                 230                 235                 240
```

```
ccg gga atc gtt cct gtt gga agc ctc cag ccg gat caa gct act tct   768
Pro Gly Ile Val Pro Val Gly Ser Leu Gln Pro Asp Gln Ala Thr Ser
                245                 250                 255

gga acc cct acc cct tct gtt agc caa agc caa tct gga cag caa tgc   816
Gly Thr Pro Thr Pro Ser Val Ser Gln Ser Gln Ser Gly Gln Gln Cys
            260                 265                 270

tgc tgc act cct cca atc aca aac cct gta atg cca act cct atg ggt   864
Cys Cys Thr Pro Pro Ile Thr Asn Pro Val Met Pro Thr Pro Met Gly
        275                 280                 285

atc agc agt aat ggg tat ccc agc tca act gcg tac gcc cca acc ctt   912
Ile Ser Ser Asn Gly Tyr Pro Ser Ser Thr Ala Tyr Ala Pro Thr Leu
    290                 295                 300

gga caa ttg gga cct tgc atc gac aca cag aag tca aca tca tcc tgc   960
Gly Gln Leu Gly Pro Cys Ile Asp Thr Gln Lys Ser Thr Ser Ser Cys
305                 310                 315                 320

gaa cca aaa gaa aag cct gta gca cag tat gga atg gaa gca tgc gct   1008
Glu Pro Lys Glu Lys Pro Val Ala Gln Tyr Gly Met Glu Ala Cys Ala
                325                 330                 335

gca cca act cca act gct gtt cta gga aat gct gag tat ctc ctt agc   1056
Ala Pro Thr Pro Thr Ala Val Leu Gly Asn Ala Glu Tyr Leu Leu Ser
            340                 345                 350

ccg ggg atg tat aat tca ctc aac tct cca tgc aac gct tgc tgc caa   1104
Pro Gly Met Tyr Asn Ser Leu Asn Ser Pro Cys Asn Ala Cys Cys Gln
        355                 360                 365

caa caa tgc tag                                                  1116
Gln Gln Cys
    370
```

<210> 4
<211> 828
<212> DNA
<213> Encephalitozoon intestinalis

<220>
<221> CDS
<222> (1)..(825)

<400> 4

```
atg ttg tta ctt ctc tca gca gtt gct ttt gtt agc gct aca gca gtc   48
Met Leu Leu Leu Leu Ser Ala Val Ala Phe Val Ser Ala Thr Ala Val
  1               5                  10                  15

cag tca ggt gtt gtc tcc cag cct aca aca ccc att ccg att ctt cct   96
Gln Ser Gly Val Val Ser Gln Pro Thr Thr Pro Ile Pro Ile Leu Pro
             20                  25                  30

gga cag ccg atg ggg ggc atg gcc aac ggg tgc act aac aag aaa cta   144
Gly Gln Pro Met Gly Gly Met Ala Asn Gly Cys Thr Asn Lys Lys Leu
         35                  40                  45
```

```
gat ggt gtt gaa ata atg aga agg aac atg gtg gaa tgc cag aag aga   192
Asp Gly Val Glu Ile Met Arg Arg Asn Met Val Glu Cys Gln Lys Arg
     50                  55                  60

aat gca gag gca aca aaa gca atg gtt gaa agg gct aat gaa aag gct   240
Asn Ala Glu Ala Thr Lys Ala Met Val Glu Arg Ala Asn Glu Lys Ala
 65                  70                  75                  80

gta gaa aca ttc aat aag gag gtc agt aaa gga cct caa aag gaa agc   288
Val Glu Thr Phe Asn Lys Glu Val Ser Lys Gly Pro Gln Lys Glu Ser
                85                  90                  95

ggc cag tgc ata gaa aaa gct gta cag ggc acc gac aga tgt att ctt   336
Gly Gln Cys Ile Glu Lys Ala Val Gln Gly Thr Asp Arg Cys Ile Leu
            100                 105                 110

gca gga ata att gat aag gct gtg aac aag cgt aag tac aga atc tcg   384
Ala Gly Ile Ile Asp Lys Ala Val Asn Lys Arg Lys Tyr Arg Ile Ser
        115                 120                 125

gat gtg gag aat agc acc tcg ctc tat aga ggc gac aaa cta att gct   432
Asp Val Glu Asn Ser Thr Ser Leu Tyr Arg Gly Asp Lys Leu Ile Ala
    130                 135                 140

cta att gtc aat gtt gac tat gga ctt cag cca att atc aaa cca aag   480
Leu Ile Val Asn Val Asp Tyr Gly Leu Gln Pro Ile Ile Lys Pro Lys
145                 150                 155                 160

aag aag aaa tcc aag ata atg gca aat ctt cct caa cca aag aga gag   528
Lys Lys Lys Ser Lys Ile Met Ala Asn Leu Pro Gln Pro Lys Arg Glu
                165                 170                 175

atg tat ttc aac cag atc gga cag ctt gtt gga gca aag gga aca ttc   576
Met Tyr Phe Asn Gln Ile Gly Gln Leu Val Gly Ala Lys Gly Thr Phe
            180                 185                 190

cct caa gac aac aag gat gaa tgc aag cca tgc gaa cct aag aag act   624
Pro Gln Asp Asn Lys Asp Glu Cys Lys Pro Cys Glu Pro Lys Lys Thr
        195                 200                 205

gtt gaa act gct tct gaa aga tgt aat ctt ggg tgc gag ctt aag gga   672
Val Glu Thr Ala Ser Glu Arg Cys Asn Leu Gly Cys Glu Leu Lys Gly
    210                 215                 220

acc tca gcc ctg ata agt aag gcc ata caa aag aag gag atc aag gag   720
Thr Ser Ala Leu Ile Ser Lys Ala Ile Gln Lys Lys Glu Ile Lys Glu
225                 230                 235                 240

agc cca aag gag ggg gac aga aac aca acc cag gaa tat gat ggt gag   768
Ser Pro Lys Glu Gly Asp Arg Asn Thr Thr Gln Glu Tyr Asp Gly Glu
                245                 250                 255

ggc tct gct gaa gat gct gaa ggc caa caa cct tct gca gac ggc gaa   816
Gly Ser Ala Glu Asp Ala Glu Gly Gln Gln Pro Ser Ala Asp Gly Glu
            260                 265                 270
```

```
ggt cta gag taa                                                        828
Gly Leu Glu
        275


<210> 5
<211> 819
<212> DNA
<213> Encephalitozoon hellem

<220>
<221> CDS
<222> (1)..(816)

<400> 5
atg ttg tta ctt ttc acc gta gtt act ctt gtt agc gct gca cag gtg    48
Met Leu Leu Leu Phe Thr Val Val Thr Leu Val Ser Ala Ala Gln Val
  1               5                  10                  15

gca cct gta act ccg cag gca gct gta cct aca caa ttc ctt cct ggt    96
Ala Pro Val Thr Pro Gln Ala Ala Val Pro Thr Gln Phe Leu Pro Gly
             20                  25                  30

gcc cag caa aag att ggc ggt gtg gac aac aga tgt gcc aac aag caa   144
Ala Gln Gln Lys Ile Gly Gly Val Asp Asn Arg Cys Ala Asn Lys Gln
         35                  40                  45

gta gaa ggt gtt caa ata ttt caa gga gac atg gcc gat tgc ccg aaa   192
Val Glu Gly Val Gln Ile Phe Gln Gly Asp Met Ala Asp Cys Pro Lys
     50                  55                  60

aga aac tcc gag gct gca aat gca atg gtt caa aga gcc aag caa aag   240
Arg Asn Ser Glu Ala Ala Asn Ala Met Val Gln Arg Ala Lys Gln Lys
 65                  70                  75                  80

gct tta gaa atc tac aat aag gag att agc aag ggc ccc aca cca aag   288
Ala Leu Glu Ile Tyr Asn Lys Glu Ile Ser Lys Gly Pro Thr Pro Lys
                 85                  90                  95

gat agc ggc cag tgc ata gaa aga gct gta caa ggt act gac agg tgt   336
Asp Ser Gly Gln Cys Ile Glu Arg Ala Val Gln Gly Thr Asp Arg Cys
            100                 105                 110

att ctt gca aaa ata atc gac aag gct gtg aac atg ctt aag tac aga   384
Ile Leu Ala Lys Ile Ile Asp Lys Ala Val Asn Met Leu Lys Tyr Arg
        115                 120                 125

atc tca aag gta gga aat gct aca gca ctc ttc aga gga aac aag cta   432
Ile Ser Lys Val Gly Asn Ala Thr Ala Leu Phe Arg Gly Asn Lys Leu
    130                 135                 140

att tct cta att ctt aat gtt gat tat gga ctt aag cca ttc ttt act   480
Ile Ser Leu Ile Leu Asn Val Asp Tyr Gly Leu Lys Pro Phe Phe Thr
145                 150                 155                 160

gtt gta aag aag aaa aca aag aga gtg ttc ccc caa ggg gat gag ctg   528
Val Val Lys Lys Lys Thr Lys Arg Val Phe Pro Gln Gly Asp Glu Leu
                165                 170                 175
```

```
aac ttc aat gga att ggt cag ctt ata gga gta aaa ggc aca ttc cct   576
Asn Phe Asn Gly Ile Gly Gln Leu Ile Gly Val Lys Gly Thr Phe Pro
            180                 185                 190

caa gac aat aat gat gaa tgc aag ccg tgt gac tct cca aag aag act   624
Gln Asp Asn Asn Asp Glu Cys Lys Pro Cys Asp Ser Pro Lys Lys Thr
        195                 200                 205

gtt gag act gtt gct gag gaa tgt aat ctt ggg tgc cag ctt aag ggg   672
Val Glu Thr Val Ala Glu Glu Cys Asn Leu Gly Cys Gln Leu Lys Gly
    210                 215                 220

acg cct ggg ttg ata agc aga gcc ata caa aag aag gag gtc aag gaa   720
Thr Pro Gly Leu Ile Ser Arg Ala Ile Gln Lys Lys Glu Val Lys Glu
225                 230                 235                 240

agc tca aag gac gga gaa aaa agc tca acc cag aac ggc gaa ggc acc   768
Ser Ser Lys Asp Gly Glu Lys Ser Ser Thr Gln Asn Gly Glu Gly Thr
                245                 250                 255

acc gat gat gaa gat gga cag caa tct ccg gac ggt aat gga cca gag   816
Thr Asp Asp Glu Asp Gly Gln Gln Ser Pro Asp Gly Asn Gly Pro Glu
            260                 265                 270

taa                                                              819
```

<210> 6
<211> 395
<212> PRT
<213> Encephalitozoon cuniculi

<400> 6

```
Met Lys Gly Ile Ser Lys Ile Leu Ser Ala Ser Ile Ala Leu Met Lys
  1               5                  10                  15

Leu Glu Asn Val Tyr Ser Ala Thr Ala Leu Cys Ser Asn Ala Tyr Gly
             20                  25                  30

Leu Thr Pro Gly Gln Gln Gly Met Ala Gln Gln Pro Ser Tyr Val Leu
         35                  40                  45

Ile Pro Ser Thr Pro Gly Thr Ile Ala Asn Cys Ala Ser Gly Ser Gln
     50                  55                  60

Asp Thr Tyr Ser Pro Ser Pro Ala Ala Pro Thr Ser Pro Val Thr Pro
 65                  70                  75                  80

Gly Lys Thr Ser Glu Asn Glu Thr Ser Pro Ser Ala Pro Ala Glu Asp
                 85                  90                  95

Val Gly Thr Cys Lys Ile Ala Val Leu Lys His Cys Asp Ala Pro Gly
            100                 105                 110

Thr Thr Ser Gly Thr Thr Pro Gly Ser Gly Pro Cys Glu Thr Pro Glu
        115                 120                 125
```

```
Gln Gln Gln Pro Leu Ser Val Ile Ser Thr Thr Pro Ala Val Pro Val
    130                 135                 140

Thr Val Glu Ser Ala Gln Ser Pro Ser Val Val Pro Val Val Pro Val
145                 150                 155                 160

Val Ala His His Gln Ala Val Pro Gly Tyr Tyr Asn Asn Gly Thr Ser
                165                 170                 175

Gly Ile Pro Gly Gln Gln Gln Ile Leu Ser Gly Thr Leu Pro Pro Gly
            180                 185                 190

Ala Thr Leu Cys Gln Gly Gln Ala Met Pro Ser Thr Pro Gly Gln Gln
        195                 200                 205

Gln Ile Leu Ser Gly Thr Leu Pro Pro Gly Val Thr Leu Cys Gln Gly
    210                 215                 220

Gln Ala Thr Pro Ser Thr Pro Gly Gln Gln Gln Val Leu Ser Gly Thr
225                 230                 235                 240

Leu Pro Pro Gly Val Thr Leu Cys Gln Gly Gln Ala Thr Pro Ser Thr
                245                 250                 255

Pro Gly Gln Gln Gln Val Leu Ser Gly Thr Leu Leu Pro Gly Ala Thr
            260                 265                 270

Leu Cys Gln Asp Gln Gly Met Pro Gly Thr Ser Gly Val Pro Gly Gln
        275                 280                 285

Gln Gly Gln Ser Ser Gly Gln Cys Cys Ala Pro Gln Ile Pro Asn Pro
    290                 295                 300

Val Met Pro Pro Ser Met Asn Ile Ser Gly Asn Gly Tyr Pro Ser Ser
305                 310                 315                 320

Thr Ala Tyr Ser Pro Asn Leu Gly Ser Leu Gly Ser Cys Val Asp Ile
                325                 330                 335

Gln Lys Thr Gly Gly Thr Ser Cys Glu Gln Lys Pro Glu Lys Ser Ala
            340                 345                 350

Thr Gln Tyr Ala Met Glu Ala Cys Ala Thr Pro Thr Pro Thr Val Ile
        355                 360                 365

Ile Gly Asn Ser Glu Tyr Leu Val Gly Pro Gly Met Tyr Asn Ala Ile
    370                 375                 380

Asn Ser Pro Cys Asn Thr Ala Val Gln Cys Cys
385                 390                 395
```

<210> 7
<211> 277
<212> PRT
<213> Encephalitozoon cuniculi

<400> 7

```
Met Leu Leu Leu Leu Ala Ile Thr Ala Val Val Ser Ala Thr Met Val
  1               5                  10                  15

His Pro Ser Ala Val Val Pro Gln Pro Ala Ala Pro Leu His Val Val
            20                  25                  30

Pro Pro Gln Gln Gln Met Gly Met Val Asn Gly Cys Thr Ser Lys Lys
        35                  40                  45

Leu Glu Gly Ala Glu Ile Met Arg Arg Asn Met Ile Glu Cys Gln Lys
    50                  55                  60

Arg Ser Ser Glu Ala Thr Lys Ala Met Ile Glu Arg Ala Asn Glu Lys
65                  70                  75                  80

Ala Val Glu Ser Phe Asn Lys Glu Val Ser Lys Gly Pro Ser Gln Lys
                85                  90                  95

Asp Gly Gly Gln Cys Ile Glu Lys Ala Val Gln Gly Thr Asp Arg Cys
            100                 105                 110

Ile Leu Ala Gly Ile Ile Asp Lys Ala Val Asn Lys Arg Lys Tyr Arg
        115                 120                 125

Ile Ser Asp Val Glu Asn Ser Thr Ser Leu Tyr Arg Gly Asp Lys Leu
    130                 135                 140

Ile Ala Leu Ile Val Asn Val Asp Tyr Gly Leu Gln Pro Ile Thr Lys
145                 150                 155                 160

Pro Lys Lys Lys Lys Ser Lys Ile Met Ala Asn Leu Pro Gln Pro Lys
                165                 170                 175

Arg Glu Met Tyr Phe Asn Gln Ile Gly Gln Leu Val Gly Ala Arg Gly
            180                 185                 190

Thr Phe Pro Gln Glu Asn Lys Glu Asp Cys Lys Pro Cys Glu Gly Pro
        195                 200                 205

Lys Lys Thr Val Glu Thr Thr Ser Glu Lys Cys Asn Leu Gly Cys Glu
    210                 215                 220

Leu Lys Gly Thr Ser Ala Leu Ile Ser Lys Ala Ile Gln Lys Lys Glu
225                 230                 235                 240

Val Lys Asp Thr Lys Glu Gly Glu Lys Ser Ala Ser Gln Asp Ser Asp
                245                 250                 255

Gly Glu Gly Thr Ala Glu Asp Ala Glu Val Gln Gln Pro Ser Ala Asp
            260                 265                 270

Gly Glu Gly Leu Glu
        275
```

<210> 8
<211> 371

<212> PRT
<213> Encephalitozoon intestinalis

<400> 8

```
Met Lys Gly Ile Ser Lys Val Leu Ser Ala Ser Ile Val Leu Met Lys
  1               5                  10                  15

Leu Lys Gly Val Tyr Ser Thr Thr Val Leu Cys Gly Asp Ser Thr Gln
            20                  25                  30

Gly Leu Gln Gly Thr Thr Gln Pro Ser Tyr Val Leu Val Pro Ser Ala
         35                  40                  45

Pro Glu Thr Ile Ala Asn Cys Gly Tyr Ser Pro Gln Asn Met Tyr Val
     50                  55                  60

Pro Ser Thr Pro Thr Thr Met Pro Ser Thr Val Pro Gly Thr Thr Gly
 65                  70                  75                  80

Glu Ser Glu Thr Pro Thr Ser Pro Thr Ser Ser Pro Thr Glu Asp Val
                85                  90                  95

Gly Thr Cys Lys Ile Ala Val Val Lys His Cys Asp Ala Pro Gly Thr
            100                 105                 110

Ser Ser Thr Pro Cys Glu Pro Glu Gln Thr Leu Ala Pro Ser Gln Pro
        115                 120                 125

Val Ala Ala Thr Ile Ala Thr Pro Leu Val Val Ala Ser Val Gln Thr
    130                 135                 140

Pro Gln Ala Ala Val Thr Ile Leu Thr Pro Lys Ala Val Ser Ala Gln
145                 150                 155                 160

Pro Ala Thr Ile Ile Ser Pro Phe Asn Gln Ala Pro Gly Tyr Tyr Asn
                165                 170                 175

Ser Ala Ile Pro Gly Gln Ile Leu Thr Gly Asn Val Leu Ser Pro Ser
            180                 185                 190

Ala Ser Ser Cys Gln Val Val Pro Gly Thr Thr Gly Ser Ser Thr Pro
        195                 200                 205

Gln Gln Leu Pro Gly Ala Val Ser Ser Gly Thr Ile Pro Cys Gln Ile
    210                 215                 220

Val Gln Gly Thr Gln Ser Ser Gly Asn Thr Pro Gly Gln Gln Phe Leu
225                 230                 235                 240

Pro Gly Ile Val Pro Val Gly Ser Leu Gln Pro Asp Gln Ala Thr Ser
                245                 250                 255

Gly Thr Pro Thr Pro Ser Val Ser Gln Ser Gln Ser Gly Gln Gln Cys
            260                 265                 270

Cys Cys Thr Pro Pro Ile Thr Asn Pro Val Met Pro Thr Pro Met Gly
        275                 280                 285
```

```
Ile Ser Ser Asn Gly Tyr Pro Ser Ser Thr Ala Tyr Ala Pro Thr Leu
    290                 295                 300

Gly Gln Leu Gly Pro Cys Ile Asp Thr Gln Lys Ser Thr Ser Ser Cys
305                 310                 315                 320

Glu Pro Lys Glu Lys Pro Val Ala Gln Tyr Gly Met Glu Ala Cys Ala
                325                 330                 335

Ala Pro Thr Pro Thr Ala Val Leu Gly Asn Ala Glu Tyr Leu Leu Ser
            340                 345                 350

Pro Gly Met Tyr Asn Ser Leu Asn Ser Pro Cys Asn Ala Cys Cys Gln
        355                 360                 365

Gln Gln Cys
    370
```

<210> 9
<211> 275
<212> PRT
<213> Encephalitozoon intestinalis

<400> 9

```
Met Leu Leu Leu Leu Ser Ala Val Ala Phe Val Ser Ala Thr Ala Val
  1               5                  10                  15

Gln Ser Gly Val Val Ser Gln Pro Thr Thr Pro Ile Pro Ile Leu Pro
            20                  25                  30

Gly Gln Pro Met Gly Gly Met Ala Asn Gly Cys Thr Asn Lys Lys Leu
        35                  40                  45

Asp Gly Val Glu Ile Met Arg Arg Asn Met Val Glu Cys Gln Lys Arg
    50                  55                  60

Asn Ala Glu Ala Thr Lys Ala Met Val Glu Arg Ala Asn Glu Lys Ala
65                  70                  75                  80

Val Glu Thr Phe Asn Lys Glu Val Ser Lys Gly Pro Gln Lys Glu Ser
                85                  90                  95

Gly Gln Cys Ile Glu Lys Ala Val Gln Gly Thr Asp Arg Cys Ile Leu
            100                 105                 110

Ala Gly Ile Ile Asp Lys Ala Val Asn Lys Arg Lys Tyr Arg Ile Ser
        115                 120                 125

Asp Val Glu Asn Ser Thr Ser Leu Tyr Arg Gly Asp Lys Leu Ile Ala
    130                 135                 140

Leu Ile Val Asn Val Asp Tyr Gly Leu Gln Pro Ile Ile Lys Pro Lys
145                 150                 155                 160

Lys Lys Lys Ser Lys Ile Met Ala Asn Leu Pro Gln Pro Lys Arg Glu
                165                 170                 175
```

```
Met Tyr Phe Asn Gln Ile Gly Gln Leu Val Gly Ala Lys Gly Thr Phe
            180                 185                 190

Pro Gln Asp Asn Lys Asp Glu Cys Lys Pro Cys Glu Pro Lys Lys Thr
        195                 200                 205

Val Glu Thr Ala Ser Glu Arg Cys Asn Leu Gly Cys Glu Leu Lys Gly
    210                 215                 220

Thr Ser Ala Leu Ile Ser Lys Ala Ile Gln Lys Lys Glu Ile Lys Glu
225                 230                 235                 240

Ser Pro Lys Glu Gly Asp Arg Asn Thr Thr Gln Glu Tyr Asp Gly Glu
                245                 250                 255

Gly Ser Ala Glu Asp Ala Glu Gly Gln Gln Pro Ser Ala Asp Gly Glu
            260                 265                 270

Gly Leu Glu
        275
```

<210> 10
<211> 272
<212> PRT
<213> Encephalitozoon hellem

<400> 10

```
Met Leu Leu Leu Phe Thr Val Val Thr Leu Val Ser Ala Ala Gln Val
  1               5                  10                  15

Ala Pro Val Thr Pro Gln Ala Ala Val Pro Thr Gln Phe Leu Pro Gly
             20                  25                  30

Ala Gln Gln Lys Ile Gly Gly Val Asp Asn Arg Cys Ala Asn Lys Gln
         35                  40                  45

Val Glu Gly Val Gln Ile Phe Gln Gly Asp Met Ala Asp Cys Pro Lys
     50                  55                  60

Arg Asn Ser Glu Ala Ala Asn Ala Met Val Gln Arg Ala Lys Gln Lys
 65                  70                  75                  80

Ala Leu Glu Ile Tyr Asn Lys Glu Ile Ser Lys Gly Pro Thr Pro Lys
                 85                  90                  95

Asp Ser Gly Gln Cys Ile Glu Arg Ala Val Gln Gly Thr Asp Arg Cys
            100                 105                 110

Ile Leu Ala Lys Ile Ile Asp Lys Ala Val Asn Met Leu Lys Tyr Arg
        115                 120                 125

Ile Ser Lys Val Gly Asn Ala Thr Ala Leu Phe Arg Gly Asn Lys Leu
    130                 135                 140
```

```
Ile Ser Leu Ile Leu Asn Val Asp Tyr Gly Leu Lys Pro Phe Phe Thr
145                 150                 155                 160

Val Val Lys Lys Lys Thr Lys Arg Val Phe Pro Gln Gly Asp Glu Leu
                165                 170                 175

Asn Phe Asn Gly Ile Gly Gln Leu Ile Gly Val Lys Gly Thr Phe Pro
            180                 185                 190

Gln Asp Asn Asn Asp Glu Cys Lys Pro Cys Asp Ser Pro Lys Lys Thr
        195                 200                 205

Val Glu Thr Val Ala Glu Glu Cys Asn Leu Gly Cys Gln Leu Lys Gly
    210                 215                 220

Thr Pro Gly Leu Ile Ser Arg Ala Ile Gln Lys Lys Glu Val Lys Glu
225                 230                 235                 240

Ser Ser Lys Asp Gly Glu Lys Ser Ser Thr Gln Asn Gly Glu Gly Thr
                245                 250                 255

Thr Asp Asp Glu Asp Gly Gln Gln Ser Pro Asp Gly Asn Gly Pro Glu
            260                 265                 270
```